# Static Liquidation and Risk Management: Controlling Variance and Expected Shortfall

Felipe Macías & Jorge P. Zubelli

*LAMCA*
*IMPA*

**Abstract**

In this report, we analyze the liquidation process of a given portfolio in the context of static strategies. By a static strategy we mean, one that does not take into account the flow of information on price changes during the control process. It tries to perform in some optimal sense to minimize losses. We seek for static strategies because those are typically the one used in situations of stress in the case of a fund collapse. Thus, we want to study the risk associated to the portfolio under a liquidation process more than the optimal strategy itself.

We observe that the classical models might lead to instabilities and display lack of robustness. We propose an improvement of the variance model which tackles the effects of the intra-day changes and impact price. We then present a model that reduces the conditional value at risk (CVaR) and the variance at the same time. We do this by performing a Cholesky decomposition of the covariance matrix and adding to it a Tikhonov regularization term. Although this report presents an optimal liquidation of portfolio point of view, we can apply the concepts presented herein to a problem of optimal allocation of portfolios. The present study may be of interest to risk management of central counterparties and clearing houses. In particular, it could be used for the computation of margins associated to portfolios.

*Keywords:* Liquidation Strategies, Monte-Carlo, Derivatives, Tikhonov Regularization, Price Impact, Intra-day Price.
*2008 MSC:* 91B30, 60A10, 91B70

## 1. Introduction

Collateral warranty is highly used in financial transactions to reduce credit and liquidity risk, in case one of the parties involved does not fulfill his obligations. It does not neces-

sarily have being presented as cash. There are several financial instruments that can be used as collateral, such as bonds, shares of stock, derivatives, etc. Therefore, it is essential to be able to determine the value and the risk of a portfolio as a whole in a situation that requires being liquidated in a short period, focusing on reducing capital loss.
Sometimes, a third party like a clearing house takes the position of managing the risk on a trade. For that, the participants have to deposit margins at the clearing house, which needs to monitor those margin levels to be sure they can cover losses in the case of settlement failure. In this scenario, it is also essential to be able to understand the risk of a portfolio in the event one of the parties falls in default.
The value of a portfolio can be determined by the market value (mark to market) of the assets, that could be calculated using some available models or historical data. Nevertheless, this might not be a good approximation of the real value, because, after the liquidation process, there is a good chance of not obtaining the full market value of the assets. This mismatch between the market value and the final price may be caused by, intra-day price variations, poor market, liquidity and size of the portfolio. Also, often such liquidation procedures occur during market stress events, and sizable transactions can negatively impact the market and produce further losses. Furthermore, the bigger the portfolio, the harder it is to find suitable buyers. We can also find different restrictions for each asset of the starting day of the process and other limitations that the portfolio's owner may encounter. Searching for efficient strategies to liquidate a portfolio is fundamental for determining its real value and its risk. In this report, we focus our interest in studying the risk of a portfolio under a liquidation strategy more than the strategy itself.
The standard practice in financial institutions, as far as margin calculations are concerned, is to consider static liquidation strategies. See [11, 10, 2]. For this reason, we shall concentrate on such strategies in this report.

The plan for this report goes as follows: In Section 2 we review basic definitions and concepts of portfolio management and risk measures. In Section 3 we discuss some liquidation-strategy models and in Section 4 we compare such models 4. We provide in Section 5 some illustrative examples of such models. In Section 6 we disscuss the issue of the execution price during the liquidation. In Section 7 we propose the use of Tikhonov-type regularization in order to improve the robustness of the liquidation process and interpret such proposal in financial terms.Section 8 presents some illustrative examples and comparisons.

## 2. Definitions and Basic Concepts

### 2.1. Definitions

Let us consider a portfolio $P$ with $N_a$ assets that have $m_i$ shares of an asset $i$, such that at the time $t = 0$ it has a mark to market (MtM) value of $\varphi_i^0$. This portfolio may includes derivatives as an asset $i$, in which the value of $\varphi_i$ depends on the price of the underlying asset $S_i$.
We want to find a strategy to liquidate this portfolio within $T$ days starting at day $t = 1$. The value of the portfolio at the time $t = 0$ will be $P_0 := \sum_{i=1}^{N_a} m_i \varphi_i^0$.
We define the vector $q \in [0, 1]^{N_a \times T}$ where $q_i^t$ represents the fraction of the wealth invested asset $i$ that will be liquidated at the time $t$. Also, we define the set

$$I := \{(i, t) \in \{1, ..., N_a\} \times \{1, ..., T\}\}.$$

Without loss of generality, we assume that the asset $i$ depends only on one risk factor $S_i$ that represents the price of one underlying share for the asset $i$. Assuming that we know the distributions of the price of the asset $S_i^t$ for all time $t \in \{1, \dots, T\}$, we define $\psi_i^t(S_i^t) := m_i(\varphi_i^t(S_i^t)e^{-rt} - \varphi_i^0)$, where $\psi_i^t$ is the present value of loss or gain of the asset $i$ at the time $t$. We are not making any assumption on the random variables $S_i^t$ other than it is a real random variable and belongs to some probability space $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.
Besides the restrictions of liquidating by the time $T$, we have other restrictions like the amount of the asset we can liquidate per day and the day we can start the liquidation.
Under these hypotheses, we define the set:

$$Q := \left\{ q \in [0, 1]^{N_a \times T} \quad | \quad q_i^t \leq k_i^t, \quad \forall (i, t) \in I, \quad \sum_{t=1}^{T} q_i^t = 1, \quad \forall i \in \{1, ..., N_a\} \right\},$$

where $Q$ is a linear bounded set and

$$\sum_{t=1}^{T} k_i^t \geq 1 \quad \forall i \in \{1, ..., N_a\}.$$

**Observation 2.1.** *We use a matrix notation for the vectors $\psi_i^t$ and $q_i^t$ where $(\cdot)_i^t$ is a reshaping of the vector $(\cdot)_{t+(i-1)T}$.*

The loss or gain obtained by liquidating the portfolio using strategy $q$ is represented by $\sum_{(i,t) \in I} q_i^t \psi_i^t$. So, given a multivariate random variable $S \in \Pi_{(i,t)\in I} \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ we define the functional:

$$Q \longrightarrow \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}) \qquad\qquad q \longrightarrow M_\psi(q) := \sum_{(i,t)\in I} q_i^t \psi_i^t(S_i^t) = \langle q, \psi \rangle, \tag{1}$$

where $M_\psi(q)$ represents the loss-gain value of the portfolio liquidated using strategy $q$. Whenever we choose $N_s$ samples of the multivariate random variable $\psi$, we shall write $\psi_k$ or $\psi_i^t(k)$ for every $k \in \{1, \ldots, N_s\}$.

### *2.2. The Liquidation Problem*

We wish to study the risk of a portfolio under a liquidation process. Every portfolio has its own risk derived from the combination of its assets, however, in the process of clearance of a portfolio, there are other risks involved. The fact that we cannot liquidate instantly at the time zero will produce a temporal exposition of the portfolio. Also, buying and selling at the initial time can cause impacts on the prices, consequently increasing the cost. Moreover, if the portfolio has a derivative, it might include special rules for selling and buying. All of these factors may result in the selection of an unsuitable strategy for liquidation, causing unnecessary exposure and expenses.

The first action when choosing a strategy is to define what type of risk the holder wants to avoid. As will be seen in the course of this report, there are different risk concepts, and each one can provide us with extremely different strategies. In the literature, there are several models that have been used to find an optimal allocation subject to reducing the defined risk [15]. These models can be easily be reformulated to reflect strategies of liquidation instead of allocation. In this section, we will review two of the classical formulations applied to our particular problem: the Variance Model and the Expected Shortfall Model.

We concentrate on static strategies rather than dynamic strategies of liquidation as our focus is to understand the risk of the portfolio in the event of mandatory liquidation, see [11, 10, 2].

### *2.3. Optimization Problem*

Suppose that we set up our definition of risk, our problem is to find how to liquidate our portfolio $P$ optimally to be defined below. For that, we define an optimization functional $F$ which is not necessarily a risk measure. Since we know the distribution of $S$ and hence the distribution of $\psi$, we can define a stochastic control problem:

$$\min_{q\in Q} F(M_\psi(q)) \tag{2}$$

## 3. Liquidation Strategy Models

In this section, we present three models: the Simpleminded Model, the Variance Model and the Expected Short all of the Loss Model. The first is a naive model and the other two have been used in allocation theory (see [7]) and that we will adapt for our purpose. Notice that we are focusing on the risk of losing not on increasing the expected returns, so we will not consider in the objective function or the restrictions the expected returns.

### 3.1. Simpleminded Model

A simple, naive strategy is to try to liquidate all the assets independently and to think that the faster you sell or buy them, the better it is.
Let us write this strategy as an optimization problem:

$$\max_{q\in Q}\sum_{i=1}^{N_a}\sum_{t=1}^{T}(T+1-t)q_i^t \tag{3}$$

Note that for this strategy, we do not consider the value of $M_\psi(q)$. This strategy is very elementary, and it does not need to be written as an optimization problem. This model is based on the principle that liquidating your asset faster you would reduce the exposition and supposedly reduce the risk of losing. However, we are going to see that this idea is not necessarily a good one, especially if every asset has different rules for selling (starting day, the maximum amount per day) and there is hedging or correlation between them.

### 3.2. Variance Model

The standard deviation is a deviation risk measure (see [20]). A commonly used model is the approach presented by Markovitz (1952) in [24] that consist of finding an optimal liquidation strategy for a portfolio minimizing the variance:

$$\min_{q\in Q}\sigma^2(M_\psi(q)) = \min_{q\in Q}\langle q, \Sigma q\rangle\,. \tag{4}$$

Here, $\Sigma$ is the covariance matrix of $\psi$ and as we said before, we eliminate the expected value constraint of the classical model. The Problem 4 is a quadratic problem in a convex and compact setting, so the problem has always a solution, which is not necessarily unique. An advantage and feature of this model is that the calculus of the covariance matrix $\Sigma$ can be very accurate and the quantity of scenarios that we take to solve the Variance Model will not affect the performance of the optimization. In practice, we can simulate a considerable

number of samples and calculate the covariance matrix with them (the calculus is just a process of matrices multiplication). This attribute of the model is advantageous for two reasons. The first is that, even if Equation 4 defines a quadratic model, it is in practice a fast model to solve. The second reason is that this model is very stable and robust.
One of the biggest disadvantages is that, since it reduces the variance on both directions, loss and gain are treated in the same way. Also it does not handle the extreme loss. Furthermore, if the problem becomes degenerate several numerical issues may arise.

### *3.3. Expected Shortfall of the Loss Model*

The Expected Shortfall Measure (also called the Conditional Value at Risk, Average Value at Risk or Expected Tail Loss) is a coherent risk measure (See [28, 29, 27]). It can be interpreted as the expected loss of a given $\alpha$-quantile. A formal definition, given a random variable $X \in L^p$ and $\alpha \in (0,1)$

$$ES_\alpha(X) := \int_0^\alpha VaR_\gamma(X) d\gamma,$$

where $VaR_\gamma(X)$ is the Value at Risk of $X$ with confidence level $\gamma$.
The fact that the Expected Shortfall is a coherent measure implies that it is a convex function of $q \in Q$. Rockafellar(2000) [23] presented a linear formulation that seeks the optimal allocation that minimizes the Conditional Value at Risk and necessarily reduces the Value at Risk.
In the article [23], they introduced the functional

$$F_\beta(q,v) = v + \frac{1}{1-\beta} \int_{y \in \mathbb{R}^{TxN_a}} \left(-M_\psi(q) - v\right)^+ p(y) dy,$$

where the solution $(q^*, v^*)$ of

$$\min_{(q,v) \in Q \times \mathbb{R}} F_\beta(q,v) \tag{5}$$

is such that $v^* = VaR_\beta(-M(q^*,\psi))$ and $F_\beta(q^*,v^*) = ES_\beta(-M(q^*,\psi))$. Instead of solving the Problem 5, they take $N_s$ sample of $\psi$ and approximate $F_\beta$ as

$$F_\beta(q,v) \approx \hat{F}_\beta(q,v) := v + \frac{1}{N_s(1-\beta)} \sum_{k=1}^{N_s} \left(-M(q,\psi_k) - v\right)^+$$

and solve

$$\min_{(q,v)\in Q\times\mathbb{R}} \hat{F}_\beta(q,v). \tag{6}$$

Using some auxiliary variables, we get the following linear problem:

$$\begin{aligned}
\min_{(q,\alpha,v,\lambda)} \quad & v + \frac{1}{N_s(1-\beta)}\sum_{k=1}^{N_s} u_k \\
\text{s. t.} \quad & q \in Q, \quad v \in \mathbb{R}, \\
& -M_{\psi_k}(q) - v \leq u_k, \quad \forall k \in \{1,\dots,N_s\}, \\
& u_k \geq 0, \quad \forall k \in \{1,\dots,N_s\}.
\end{aligned} \tag{7}$$

As mentioned before, ideally we are looking for strategies that reduce the risk of loss, in other words just the risk of the negative part of $M_\psi(q)$. So, in Cont (2013)(see [18]) a loss-based risk measure was introduced where they consider measures that focus only on the loss part. The conditional value of the loss is an example of this measure. Hence, we can easily extend the linear model in [23] for the Expected Shortfall to a model for the Expected Shortfall of the Loss (ESL). Remembering that $M_\psi(q)^- = \max\{-M_\psi(q), 0\}$,

$$\hat{F}_\beta(q,v) := v + \frac{1}{N_s(1-\beta)}\sum_{k=1}^{N_s}\left(M_{\psi_k}(q)^- - v\right)^+.$$

The problem can be written as a linear programming introducing a new auxiliary variable $\gamma \in \{1,\dots,N_s\}$

$$\begin{aligned}
\min_{(q,\alpha,v,\lambda)} \quad & v + \frac{1}{N_s(1-\beta)}\sum_{k=1}^{N_s} u_k, \\
\text{s. t.} \quad & q \in Q, \quad v \in \mathbb{R}, \\
& \gamma_k - v \leq u_k, \quad \forall k \in \{1,\dots,N_s\}, \\
& \gamma_k \geq -M_{\psi_k}(q), \quad \forall k \in \{1,\dots,N_s\}, \\
& u_k \geq 0, \quad \gamma_k \geq 0, \quad \forall k \in \{1,\dots,N_s\}.
\end{aligned} \tag{8}$$

Among the advantages, the ESL is a convenient representation of risks, as it measures the downside risk. Also, it is applicable to non-symmetric loss distribution. Another properties are that is a convex model respect the portfolio position and it is a loss-based risk measure which can be used a linear programming approach to solve the minimization.
Among the disadvantages, we can cite that the size of the LP increases when we increase

the number of scenarios and it focuses only on the tail of the loss, do not reduce more general risk as the expected of the loss.

## 4. Model Comparison

Choosing one model over the other will depend on the investor profile. However, it is important to make a comparison between them to have a better understanding of their properties. The first thing we are going to compare is the computational cost of each model. Second, as we are working with risk estimates for solving the optimization, we need to evaluate how robust are these estimations. So, we will define two measures that will help us to study the stability of the models reviewed.

### *4.1. Computational efficiency*

Let us write $R_q := N_a(1 + 2T)$, Table 4.1 shows the numbers of the constraints and variables for each model. A first observation is that the Simpleminded Model has the same number of constraints and variables than the Variance model, besides that fact that one is linear and the other quadratic, the difference comes from the fact that the Simpleminded Model does not have any information about the behavior of the portfolio. In the Variance Model, although the model is non-linear the fact that the optimization does not depend on the amount of scenarios (the input is just the covariance matrix), makes this model faster than the ESL Models when the number of simulations is big. Because the ESL Model depends on the number of scenarios, we need to choose an appropriate set of scenarios that are representative of the distribution $\psi$ but not so big that the optimization algorithm could not converge in a reasonable time.

| *Model* | *Variables* | *Constraints* | *Problem Type* |
|---|---|---|---|
| Simpleminded | $TN_a$ | $R_q$ | Linear |
| Variance | $TN_a$ | $R_q$ | Quadratic Convex |
| Ex. Sh. Loss | $TN_a + 1 + 2N_s$ | $R_q + 2N_s$ | Linear |

**Observation 4.1.** *We are going to refer to the order of the optimization problem as $O(n)$, which means that the computational time of the model depends linearly of $n$. There, we write $O(N_s)$ when the optimization problem depends linearly on the size of the sample, and we write $O(N_aT)$ when it depends on the number of asset and time step $T$. When it depends on the size of the sample $Ns$ and the number of asset and time step $N_aT$ independently, we write only $O(N_s)$ because in practice $N_s \gg N_aT$.*

*4.2. Robustness*

In practice, we can not work directly with the process $\psi$ because of its complexity and because we do not want to make any assumption about it. Because every process could be very complex independently and combine, we are going to use Monte-Carlo simulations to compute an approximation of the risk measures.
Consider that we have a risk functional $F$ and $q^* \in Q$ solution of Equation 2 and suppose that $\Omega$ is a huge finite set of all possible scenarios of $\psi$. To solve Problem 2, we use the linear models taking a sample with $K$ scenarios of $\Omega$ and using an approximation $\hat{F}$ of $F$ to solve

$$\min_{q \in Q} \hat{F}(\{M_{\psi_k}(q)\}_{k \in K}). \tag{9}$$

We are going to denote $q^K$ the solution of Problem 9. We want to study how close is $M_\psi(q^K)$ from $M_\psi(q^*)$.
There is a trade-off between the computational time and the quality of the solution if we increase the number of scenarios we are going to have a better solution but with an efficiency cost. We need to study the size of a representative number of scenarios. For this reason, we are going to use two tests for the robustness. They are going to be called a Risk Functional (RF) Robustness and Cumulative Distribution Function (CDF) Robustness. The RF Robustness concerns how the risk value used in the optimization is behaving when the numbers of scenarios are increasing. The CFR studies the $L^\infty$ norm of the cumulative function $C_F$ of $M_\psi(q)$.

**Definition 4.1.** *The Risk Functional Robustness is the study of the behavior of*

$$\frac{|F(M_\psi(q^K)) - F(M_\psi(q^{K'}))|}{F(M_\psi(q^K))}, \tag{10}$$

*when* $|K|, |K'| \to |\Omega|$.

**Definition 4.2.** *The Cumulative Distribution Function Robustness is the study of the behavior of*

$$\|C_F(M_\psi(q^K)) - C_F(M_\psi(q^{K'}))\|_\infty, \tag{11}$$

*when* $|K|, |K'| \to |\Omega|$.

## 5. Illustrative Example

Let us consider a a simple and fictional portfolio formed by
We simulated $1,000,000$ scenarios for the share $S$ using a Brownian motion (see [26]) with

| *Asset* | *Product* | *Position* | *Exp Day* | *Strike* | *Max p/ Day* | *Initial Day* |
|---|---|---|---|---|---|---|
| 1 | Option Call | -4000 | 60 | 1.2 | 2000 | 5 |
| 2 | Option Put | 1000 | 60 | 1.2 | 1000 | 5 |
| 3 | Forward | 1000 | 60 | 1.2 | 1000 | 2 |
| 4 | Share | 1000 | N/A | N/A | 1000 | 1 |

Table 1: Portfolio

annual volatility of 0.25 and annual drift 9.00%. The annualized risk-free interest rate was 8.00%. The three derivatives are associated with the same share $S$ of the first asset with initial value $S_0 = 1$, and the mark-to-market values of the options were calculated using the Black-Scholes formula (see [25]). The initial value of the portfolio $P_0$ is 931.3.
The Simpleminded Model did not need the simulations, for the Variance Model, we used all the scenarios to calculate the covariance matrix, for the ESL Model, we choose $6,500$ random scenarios from the $1,000,000$ scenarios simulated. Finally, for the ESL Model we used the confidence level $\beta = 0.05$.[1]

Table 2 shows the risk statistics of the solution of the different models.

| *Models* | *Std. Des.* | $CVaR_{0.05}$ | $VaR_{0.05}$ | $\mathbb{E}(X^-)$ | *Op. Time (s)* |
|---|---|---|---|---|---|
| Simpleminded | 83.2 | 213.0 | 137.5 | 65.4 | 9.4 |
| Variance | 19.2 | 49.8 | 26.2 | 14.8 | 5.5 |
| Ex. Sh. Loss | 19.6 | 49.7 | 25.3 | 13.9 | 22.9 |

Table 2: Statistics of $M_\psi(q)$.

An important question is how the holder of the portfolio will value it in a liquidation

| *Models* | $CVaR_{0.05}$ | $VaR_{0.05}$ | $\mathbb{E}(X^-)$ |
|---|---|---|---|
| Simpleminded | 77.1 | 85.2 | 93.0 |
| Variance | 94.6 | 97.2 | 98.4 |
| Ex. Sh. Loss | 94.7 | 97.3 | 98.5 |

Table 3: Percentage of the value the initial portfolio that the holder will consider for warranty.

condition. In this example, the initial value of the portfolio is 931.3, however as we showed

[1]The optimization problems were run using Gurobi Optimizer, see [1]. We choose to use it due to its satisfactory handling of sparsity.

in Table 2 different strategies give different risks. For example, use the strategy from the Simpleminded Model and the holder has a very adverse risk profile, he should consider the $CVaR$ value, i.e., price the warranty as 77.1% of the original value, however if the holder does not have a very risk aversion he may use the percentage of the expected of the loss, which is 93.0%, see Table 3. On the other hand, if the holder consider the strategy of the Variance Model or the ESL Model, and he has a very adverse risk profile, he is going to price the warranty with a bigger value, 94.6% and 94.7% respectively, than using the strategy from the Simpleminded Model and having a small adverse risk profile. Table 3 shows us the advantage of using strategies that focus in reducing the risk, for example, the holder could ask for less collateral, or the owner will have more margin before the holder asks for more collateral.

Even more, Figures 1, 2 and 3 show the optimal strategies $q$ and the histograms of $M_\psi(q)$, where they present that the Simpleminded Model is much risky than the others. In figures 2 and 3 we can see that it does not matter if we can liquidate an asset at the initial time, is more important to reduce the general risk. Also, in these figures, we can observe that although the solutions $q$ are different for the Variance Model and the ESL Model, the risks are similar (see also Table 2 and Table 3). We will see later that these risks will differ when we incorporate the intra-day variations.

**Remark 5.1.** *Notice that the histograms and the calculus of the statistics were made using all the scenarios. It does not matter if we took a few scenarios to find $q$ for the optimization problem, as we want to study the behavior of the random variable $M_\psi(q)$.*

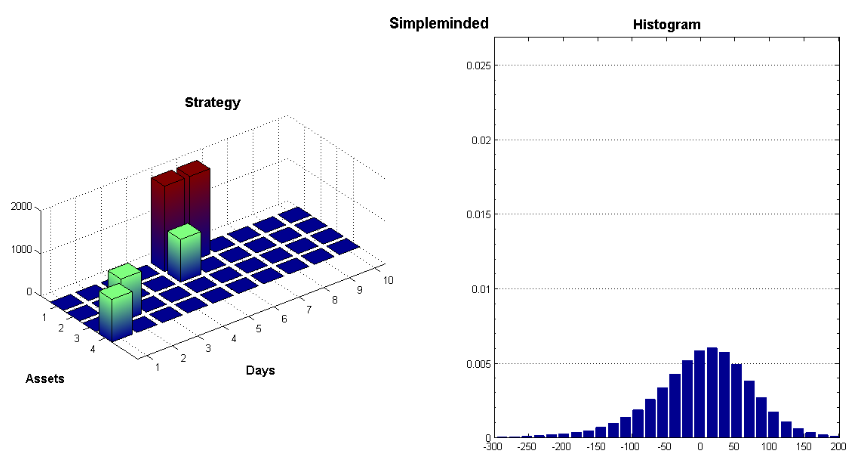


Figure 1: Strategy $q$ and distribution of $M_\psi(q)$ for the Simpleminded Model.

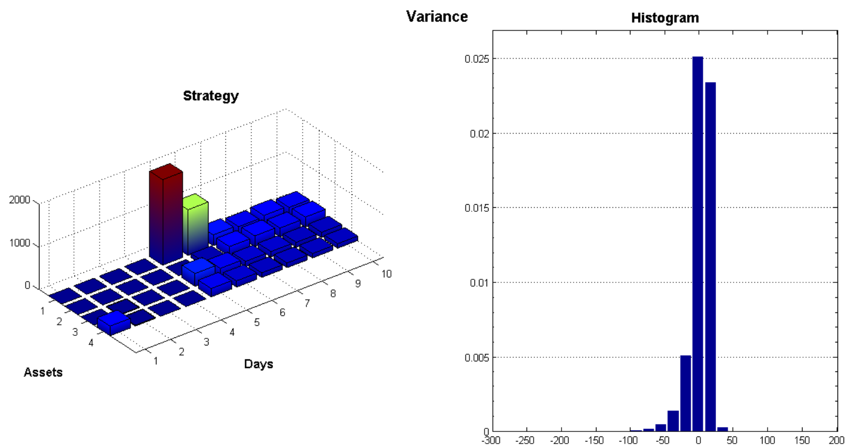


Figure 2: Strategy $q$ and distribution of $M_{\psi}(q)$ for the Variance Model.

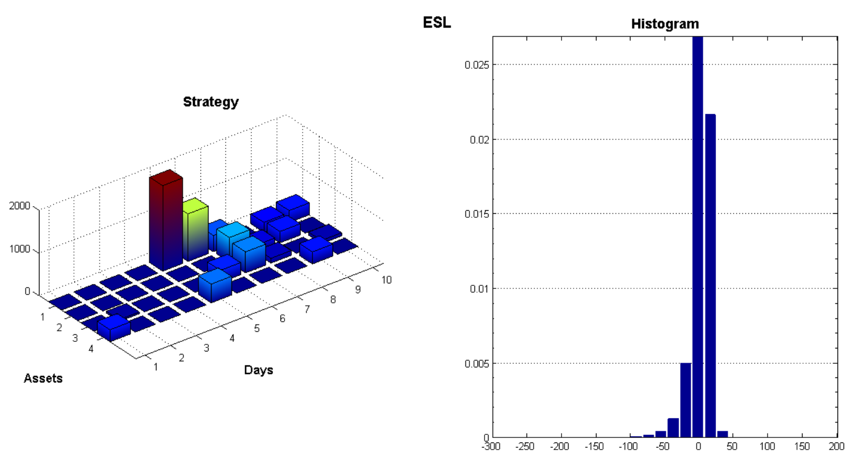


Figure 3: Strategy $q$ and distribution of $M_{\psi}(q)$ for the Expected Shortfall Model.

### *5.1. Robustness of the ESL Model*

We are going to study the behavior of the solution using the robustness defined above for the ESL model, which depends on the number of samples.

We ran the model with $1,000$ scenarios, then with other $1,500$ different scenarios and so on until $6,500$ scenarios. The graphics on the left side of Figures 4 and 5 show that the ESL Model is very robust, as it did not take more than $4,000$ scenarios to become very stable. See the graphics on the right side of Figures 4 and 5.

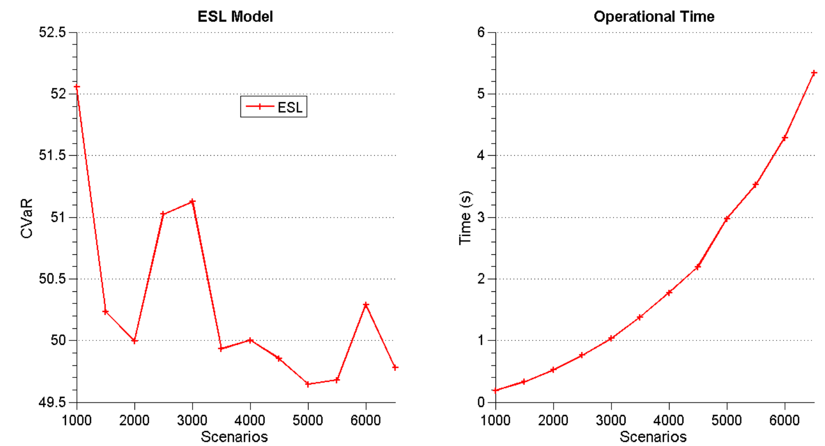


Figure 4: CDF Robustness.

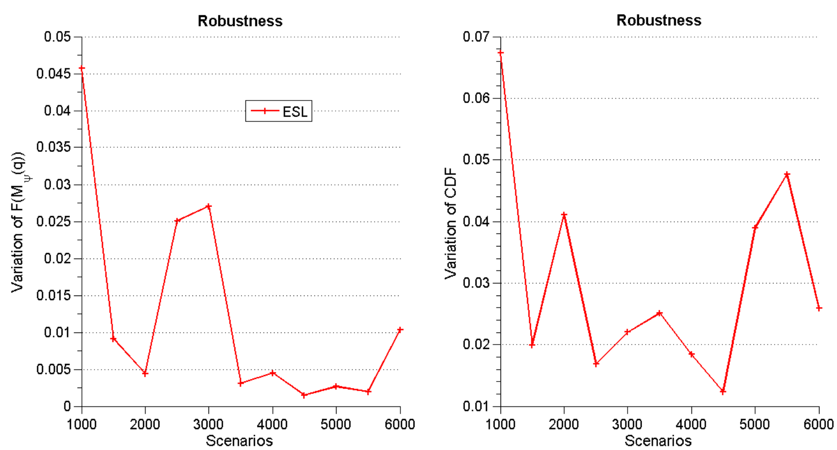


Figure 5: RF Robustness.

## 6. Execution Price

In almost every sell or buy process of an asset, the final execution price will differ from the market value. This discrepancy is due to liquidation issues in the market, the alteration of the price caused by the transaction itself or transaction costs, the bid/offer price and the lack of market depth.

In the previous section, we reviewed some essential risk aversion models to find discrete strategies. Nevertheless, the strategy $q_t^i$ can be executed at any time during the interval $(t-1, t]$, not necessarily at a fixed time $t$. Thus, this discretization causes a loss of information of the price between periods (intra-day price), besides these models do not consider the issues mentioned above. Motivated by these issues, different models have been introduced. Two models we should highlight are the articles of Bertsimas & Low

(1998) (see [17]) and Almgren & Chriss (2000) (see [12]).
We will briefly review the models in [17] and [12]. As a consequence of their work, we will present a simplified model, where this simplification makes it easier to incorporate the perturbation effect in the liquidation models.

*6.1. Basic Concepts and Models' Review*

*6.1.1. Intra-day Price*

The models we discussed in Section 2 lead to optimal discrete strategies of the amount we should liquidate every day until the day $T$. The problem with such approach is that it assumes that the liquidation of the asset $i$ will be at the end of the period or during an exact time. However in practice, the execution takes place all along caused by the lack of liquidity of the market or by trader's decision. Thus, being exposed to intra-day variations. In Figure 6, we show the effect of time discretization, for each day and only one sample, it shows we may be subject to a lot of possible prices.

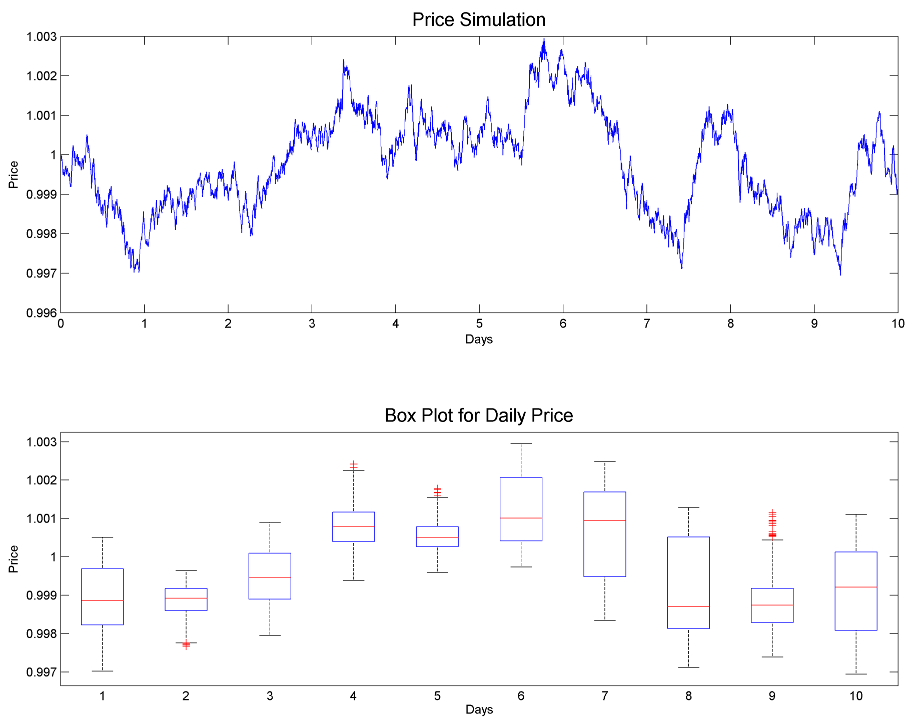


Figure 6: Up: One sample for 10 days with time steps of 30 minutes. Down: The box-plot for the daily price variation.

*6.1.2. Price Impact*

Bertsimas and Low (1998) (see [17]) present optimal dynamic strategies to minimize the expected cost of trading a large block of equities over a fixed time horizon. Moreover, they present a model for the price of an asset affected by the impact of the amount of the asset.

One such model is the one called 'linear-percentage temporary' (LPT). Another model is presented in Almgren and Chriss (2000) (see [12]), it is based on the idea of [17]. Nevertheless, it aims to find static strategies that minimize not only the expected cost but also the volatility risk. They suppose that they have a permanent impact on the price and a temporary impact on the price

*6.2. Execution Price Model*

Recall that we are seeking strategies that minimize the risk of $M_\psi(q)$, where

$$M_\psi(q) = \sum_{(i,t)\in I} \psi_i^t q_i^t$$

and

$$\psi_i^t = m_i(\varphi_i^t e^{-rt} - \varphi_i^0).$$

Hence, if we consider that we have a permanent impact over $\varphi$ that depends on $q \in Q$, $M_\psi$ would lose the linearity concerning $q \in Q$. Thus, in the minimization problems, the objective function would lose the quadratic behavior in the Variance Model and the linearity in the ESL Model. Another difficulty is to find good estimations for the parameters of the models.

The issues presented above motivate us to simplify the model, instead of supposing that we know the behavior of any impact function, we consider that the price for each asset $i$ has only temporary perturbation $\delta$. With this, we write,

$$\tilde{\varphi}_i^t := \varphi_i^t(1 + \delta_i^t), \tag{12}$$

where $\tilde{\varphi}_i^t$ is the execution price, $\varphi_i^t$ is the expected price at the time $t$ with no-impact and $\delta_i^t$ is the perturbation caused by the intra-day variations and the price impact. Notice that Equation 12 is a simplification of the LPT model presented in (see [17]) with the difference that we do not make any a priori hypothesis from the behavior of the random variable $\delta$. Also, $\delta$ could incorporate some transaction costs.

**Observation 6.1.** *Hereafter, we will refer as intra-day price both the effect of the intra-day price and the price impact.*

Therefore, we define the execution loss/gain $\tilde{\psi}$ for the asset $i$ at the time $t$:

$$\begin{aligned}\tilde{\psi}_i^t &= m_i(e^{-rt}\varphi_i^t(1+\delta_i^t) - \varphi_i^0) \\ &= \psi_i^t(1+\delta_i^t) + m_i\varphi_i^0\delta_i^t.\end{aligned} \tag{13}$$

And following the idea of [12], we are going to focus on reducing the variance of $M_{\tilde{\psi}}(q)$, changing $\psi$ by $\tilde{\psi}$. We now introduce some hypothesis on $\{\delta_i^t\}_{(i,t)\in I}$,

1. $\{\delta_i^t\}_{(i,t)\in I}$ is a set of independent random variables for all $(i,t) \in I$ and $\{\delta_i^t\}_{t=1}^T$ are identically distributed for all $i \in \{1,\ldots,N_a\}$.
2. $\delta_j^s$ is independent of $\psi_i^t$ for all $(j,s) \in I$ and $(i,t) \in I$.

We recall the following properties for two independently random variables $X$ and $Y$:

1. $\operatorname{cov}(X,Y) = 0$,
2. $\sigma^2(XY) = \sigma^2(X)\sigma^2(Y) + \mathbb{E}^2(Y)\sigma^2(X) + \mathbb{E}^2(X)\sigma^2(Y)$,
3. $\operatorname{cov}(X,XY) = \mathbb{E}(Y)\sigma^2(X)$.

With these, we calculate the variance of only $\tilde{\psi}_i^t$:

$$\begin{aligned}\sigma^2(\tilde{\psi}_i^t) &= \sigma^2(\psi_i^t(1+\delta_i^t)) + (m_i\varphi_i^0)^2\sigma^2(\delta_i^t) + 2m_i\varphi_i^0 \operatorname{cov}(\psi_i^t(1+\delta_i^t),\delta_i^t) \\ &= \sigma^2(\psi_i^t) + \sigma^2(\psi_i^t\delta_i^t) + 2\operatorname{cov}(\psi_i^t,\psi_i^t\delta_i^t) + (m_i\varphi_i^0)^2\sigma^2(\delta_i^t) + 2m_i\varphi_i^0(\operatorname{cov}(\psi_i^t,\delta_i^t) + \operatorname{cov}(\psi_i^t\delta_i^t,\delta_i^t)) \\ &= \sigma^2(\psi_i^t) + \sigma^2(\psi_i^t)\sigma^2(\delta_i^t) + \mathbb{E}^2(\psi_i^t)\sigma^2(\delta_i^t) + \mathbb{E}^2(\delta_i^t)\sigma^2(\psi_i^t) \\ &\quad + 2\mathbb{E}(\delta_i^t)\sigma^2(\psi_i^t) + (m_i\varphi_i^0)^2\sigma^2(\delta_i^t) + 2m_i\varphi_i^0\sigma^2(\delta_i^t)\mathbb{E}(\psi_i^t) \\ &= \sigma^2(\psi_i^t)(1 + 2\mathbb{E}(\delta_i^t) + \mathbb{E}^2(\delta_i^t)) + \sigma^2(\delta_i^t)(\sigma^2(\psi_i^t) + (m_i\varphi_i^0)^2 + \mathbb{E}^2(\psi_i^t) + 2\phi_i^0\mathbb{E}(\psi_i^t)) \\ &= \sigma^2(\psi_i^t)(1+\mathbb{E}(\delta_i^t))^2 + \sigma^2(\delta_i^t)(\sigma^2(\psi_i^t) + (m_i\varphi_i^0 + \mathbb{E}(\psi_i^t))^2).\end{aligned} \tag{14}$$

Using that the perturbations $\delta_i^t$ are independent, we conclude that the covariance of $\tilde{\psi}_i^t$ and $\tilde{\psi}_{i'}^{t'}$ where $(i,t) \neq (i',t')$ is

$$\operatorname{cov}(\tilde{\psi}_i^t, \tilde{\psi}_{i'}^{t'}) = \operatorname{cov}(\psi_i^t, \psi_{i'}^{t'}). \tag{15}$$

Hence we can write the covariance matrix for $\tilde{\psi}$ as

$$\tilde{\Sigma} := \Sigma(\tilde{\psi}) = \Sigma(\psi) + \Delta(\psi,\delta), \tag{16}$$

where $\Delta(\psi, \delta)$ is a diagonal matrix with $\sigma^2(\psi_i^t)((1+\mathbb{E}(\delta_i^t))^2 - 1) + \sigma^2(\delta_i^t)(\sigma^2(\psi_i^t) + (m_i\varphi_i^0 + \mathbb{E}(\psi_i^t))^2)$ in the diagonal.

Therefore, if we want to reduce the effect of the intra-day price in the liquidation process, we can solve the Variance Model considering the covariance matrix $\Sigma(\tilde{\psi})$ instead of $\Sigma(\psi)$.

$$\min_{q\in Q} \langle q, \tilde{\Sigma} q\rangle. \tag{17}$$

**Observation 6.2.** *A simple idea to incorporate the perturbation in the ESL Model is to produce samples not only for $S_i^t$ but also of $\delta_i^t$ and solve the optimization adding these samples. However, this approach significantly affects the performance of the models because the linear problems depend directly on the amount of scenarios. In the next section, we will introduce a technique to treat the intra-day price in the linear models without affecting their performances.*

**Proposition 6.1.** *Assume that $\mathbb{E}[\delta_i^t] = 0$ and $\sigma^2(\delta_i^t) = \sigma_{\delta_i}^2$, $\forall (i,t) \in I$, and that the daily limitation for liquidation is $k_i^t = 1$, $\forall (i,t) \in I$. Then the solution of*

$$\min_{q\in Q} \frac{1}{2}\langle q, \Delta(\psi, \delta) q\rangle \tag{18}$$

*is*

$$(q_i^t)^* = \frac{1}{(d^t)^2}\left(\sum_{t=1}^{T} \frac{1}{(d^t)^2}\right)^{-1},$$

*where $(d_i^t)^2 := \sigma^2(\psi_i^t) + (\phi_i^0 + \mathbb{E}(\psi_i^t))^2$ and is independently of $\sigma^2(\delta_i^t)$.*

**Proof.** First, notice that $(d_i^t)^2 > 0$, because $\sigma^2(\psi_i^t) = 0$ and

$$m_i\varphi_i^0 + \mathbb{E}(\psi_i^t) = m_i\mathbb{E}(\varphi_i^t e^{-rt}) = m_i\varphi_i^t e^{-rt} = 0$$

means that certainly the asset $i$ is going to loose everything at the time $t > 0$. Problem 18 is then equivalent to

$$\begin{aligned} \min_{q} \quad & \frac{1}{2}\sum_{(i,t)\in I} (q_i^t \sigma_{\delta_i}^2 d_i^t)^2, \\ \text{s. t.} \quad & \sum_{t=1}^{T} q_i^t = 1, \quad \forall i \in \{1, \dots, N_a\}, \\ & q_i^t \geq 0, \quad \forall (i,t) \in I. \end{aligned} \tag{19}$$

Then, if $\sigma_{\delta_i}^2$ does not depend on $t$ and the Problem 19 can be solved independently for each asset $i$, then without loss of generality we can write:

$$\begin{aligned} \min_q \quad & \frac{1}{2}\sum_{t=1}^{T}(q^t d^t)^2 \\ \text{s. t.} \quad & \sum_{t=1}^{T} q^t = 1, \quad q^t \geq 0, \quad \forall t \in \{1,\ldots,T\}. \end{aligned} \tag{20}$$

The optimization problem in Equation 20 has a strictly convex objective function in a compact space, hence it has a unique solution, and we can use the Karush-Kuhn-Tucker (KKT) conditions (see [8]) to find it. Therefore

$$q_*^t = \frac{1}{(d^t)^2}\left(\sum_{t=1}^{T}\frac{1}{(d^t)^2}\right)^{-1}. \tag{21}$$

■

**Observation 6.3.** *The model presented in this section is not far away from the model of [12]. Despite not assuming a dependence on $q$ of the price impact, we obtain that the variance of the intra-day price in our model is*

$$\sum_{t=1}^{T}(q^t)^2(d^t)^2.$$

*Therefore, we also obtain a quadratic minimization over $q$ in function of the variance.*

## 7. A Tikhonov Type Regularization to Reduce Both Risks Simultaneously

Reviewing Section 2 and 6 we can find some relevant properties of the Variance Model. In Section 2 we showed that the Variance Model is a robust model and it could be solved within at a reasonable computational time. Furthermore, in Section 6 we showed that it can incorporate intra-day risk without altering its efficiency. The idea consists in replacing the covariance matrix $\Sigma$ by $\Sigma + \Delta$ and proceed exactly as before.

On the other hand, the Expected Shortfall of the Loss (ESL) Model seems to be more effective controlling losses. However, as we studied in Section 5, these linear models present a trade-off between robustness and computational efficiency, caused by the need of numerous scenarios for the estimation of the functional to represent well uncertainty. Computational limitations preclude an arbitrary choice of a number of the sample. For the same reason, we do not recommend simulating the intra-day perturbations while generating the asset

sample as we showed in the Remark 6.2.
In this section, we shall present a formulation that takes the advantages of the VM and used to increase the robustness of the ESL Model. Furthermore, we incorporate the intra-day risk without increasing the numerical complexity. This technique will also help to control for one side the risk of large losses and on the other side to control a more global risk as is the variance.

### *7.1. Preliminaries*

Consider the linear formulation $\tilde{F}$ of the Expected Shortfall of the Loss (ESL) Model. As we remark in Observation 6.2, an idea is to incorporate the effect of the intra-day price is to minimize the linear models using $\tilde{\psi}$ instead of $\psi$, i.e.

$$\min_{q \in Q} \tilde{F}(M_{\tilde{\psi}}(q)), \tag{22}$$

where $\tilde{\psi}_i^t$ depends on the underlying share $S_i^t$ and the perturbation $\delta_i^t$ as in Equation 13. Then, to solve Problem 22, we need not only to simulate scenarios for all $S_i^t$ but also for all $\delta_i^t$. This means that for every simulation $k \in \{1, \dots, N_s\}$ we will have $N_\delta$ simulations, where $N_\delta$ represent the number of simulations of $\delta_i^t$. Hence, the linear problem becomes a model of order $O(N_s N_\delta)$. Which leads us to conclude that this idea will bring operational issues because the runtime is going to be extremely high.
Another approach is to reduce the variance of $M_{\tilde{\psi}}(q)$ at the same time we minimize $\tilde{F}(M_\psi(q))$, we do this adding $\mathrm{Var}(M_{\tilde{\psi}}(q)) = \langle q, \tilde{\Sigma} q \rangle$ to the objective function as in [4, 3]. Thus, we now consider the problem

$$\min_{q \in Q} \gamma \tilde{F}(M_\psi(q)) + (1-\gamma)\langle q, \tilde{\Sigma} q \rangle, \tag{23}$$

where $\gamma \in (0,1)$ is a parameter.
Recalling that $\tilde{\Sigma}$ is a covariance matrix we can rewrite Problem 23. Indeed, since $\tilde{\Sigma}$ is a covariance matrix, it is a symmetric and positive semi-definite matrix. Hence, we can use an eigenvector decomposition to write

$$\tilde{\Sigma} V = V D,$$

where $V$ is an orthonormal matrix with the eigenvector of $\tilde{\Sigma}$ in the columns and $D$ is diagonal, with nonnegative eigenvalues. Let $W := V D^{\frac{1}{2}}$ and use it as a Cholesky decomposition for $\tilde{\Sigma}$,

$$\tilde{\Sigma} = VDV' = VD^{\frac{1}{2}}D^{\frac{1}{2}}V' = (VD^{\frac{1}{2}})(VD^{\frac{1}{2}})' = WW'.$$

Hence,

$$\langle q, \tilde{\Sigma}q\rangle = \langle q, WW'q\rangle = \langle W'q, W'q\rangle = \|W'q\|_2^2,$$

and Problem 23 becomes

$$\min_{q\in Q} \gamma \tilde{F}(M_\psi(q)) + (1-\gamma)\|W'q\|_2^2, \tag{24}$$

where the right hand side of the term in Equation 24 is a Tikhonov regularization with term $W$(see [5]) for the Linear Model using an $L^2$ semi-norm respect to $q$.
The benefits of adding the regularization terms are several. First of all, as we will show in the next examples the model becomes more stable. Secondly, it is a simple tool for intra-day-risk control. On the other hand, one of the disadvantages of the quadratic regularization is that the models become harder to solve computationally. Table 4.1 shows that ESL Model has $O(N_s)$ variables and $O(N_s)$ constraints, hence Problem 23 is a quadratic problem with $O(N_s)$ variables and $O(N_s)$ constraints. Nevertheless, avoiding computational complexity is one of our goals. Thus this approach will not help us to keep the model simple and fast.

*7.2. The Semi-Norms $\|\cdot\|_{W^1}$ and $\|\cdot\|_{W^2}$*

We remark that the regularization term is just the semi-norm computed according to the structure defined by $W$. Then, we are going to write $\|\cdot\|_{W^2}$ to refer to this semi-norm,

$$\|q\|_{W^2} := \|W'q\|_2 = \left(\sum_{(j,s)\in I} \langle q, w_j^s\rangle^2\right)^{\frac{1}{2}}. \tag{25}$$

Remembering that $\|\cdot\|_W$ is a semi-norm if:

- $\|aq\|_W = |a|\|q\|_W$, $\forall a \in \mathbb{R}$.
- $\|q+p\|_W \leq \|q\|_W + \|p\|_W$.

Considering that the semi-norm $\|\cdot\|_{W2}^2$ is as a quadratic regularization term and on the other hand we are trying to avoid the issues of adding this term in large-scale linear problem, it is natural to ask if we can replace the $L^2$ norm of $W'q$ for an $L^1$ norm in Equation 24, i.e.,

$$\min_{q\in Q} \gamma\tilde{F}(M_\psi(q)) + (1-\gamma)\|W'q\|_1. \tag{26}$$

**Observation 7.1.** *The idea of changing the $L^2$ norm to the $L^1$ norm was motivated by the work in [16] applied to an inverse problem. The objective of this change of the norm is different from our case. However, it inspired us to take this direction.*

To study the implications of this change, let us first define

$$\|q\|_{W^1} := \|W'q\|_1 = \sum_{(j,s)\in I} |\langle q, w_j^s\rangle|. \tag{27}$$

And recall the following relationship between the norms $L^1$ and $L^2$,

$$\|q\|_{L^2} \leq \|q\|_{L^1} \leq (N_a T)^{\frac{1}{2}} \|q\|_{L^2}. \tag{28}$$

This property is easy to check using the definition of the norm, and it will help us to prove Proposition 7.1, which it is a result that assures us that the semi-norm $W^1$ will keep $q$ close to the optimal variance. In practice, the optimal variance could not be zero, so if we minimize the term $\|q\|_{W^1}$, we can not be sure that we are close to minimum variance. To avoid this problem, we first solve

$$q_a := \operatorname*{argmin}_{q\in Q} \langle q, \tilde{\Sigma} q\rangle. \tag{29}$$

And we use this as a priori term in Problem 26. Therefore, we solve

$$\min_{q\in Q} \left\{ \gamma\tilde{F}(M_\psi(q)) + (1-\gamma)\|(q-q_a)\|_{W^1} \right\}. \tag{30}$$

**Proposition 7.1.** *If $q_a \in Q$ solves the Problem 29 and $\{q_n\} \subseteq Q$ is such that $\|q_n - q_a\|_{W^1} \xrightarrow{n\to\infty} 0$, then $\mathrm{Var}(M_{\tilde{\psi}}(q_n)) \xrightarrow{n\to\infty} \mathrm{Var}(M_{\tilde{\psi}}(q_a))$.*

**Proof.** Take $q_a \in Q$ solution of Problem 29 and $q_n \in Q$, then

$$\mathrm{Var}(M_{\tilde{\psi}}(q_a)) \leq \mathrm{Var}(M_{\tilde{\psi}}(q_n)) \Rightarrow \|W'q_a\|_2^2 \leq \|W'q_n\|_2^2,$$

then

$$\|q_a\|_{W^2} \leq \|q_n\|_{W^2},$$

by the definition of semi-norm,

$$\|q_n\|_{W^2} \leq \|q_n - q_a\|_{W^2} + \|q_a\|_{W^2}$$

and using the inequality for the norms in Equation 28

$$\|q_n - q_a\|_{W^2} \leq \|q_n - q_a\|_{W^1}.$$

Therefore,

$$\|q_a\|_{W^2} \leq \|q_n\|_{W^2} \leq \|q_n - q_a\|_{W^1} + \|q_a\|_{W^2}. \tag{31}$$

Letting $\|q_n - q_a\|_{W^1} \xrightarrow{n\to\infty} 0$ Equation 31, then $\|q_n\|_{W^2} \xrightarrow{n\to\infty} \|q_a\|_{W^2}$. Finally, remembering that $\|q\|^2_{W^2} = \langle q, \tilde{\Sigma} q\rangle = \mathrm{Var}(M_{\tilde{\psi}}(q))$ we conclude the result. ■

**Remark 7.1.** *The last result shows the importance of using the a priori $q_a$. Without it, we could not be close to the optimal variance whenever we add the semi-norm $W^1$ to the linear model.*

**Remark 7.2.** *The used of an $L^1$ norm is not to obtain a spare solution of $q$. In fact, because we are using an a priori solution $q_a$ that comes from a quadratic minimization we cannot guarantee a spare solution. The change of norm is to write the minimization problem as a linear programming problem as we will show in Lemma 7.1.*

*7.3. Linearization of $\|\cdot\|_{W^1}$*

Proposition 7.1 shows that if we seek strategies $q \in Q$ close in the semi-norm $\|\cdot\|_{W1}$ to a solution $q_a$ of Problem 29, the variance of $M_{\tilde{\psi}}(q)$ will be near the optimal. Moreover, the change of the norm leads us to an equivalent linear model for Problem 30, as we prove in the following lemma.

**Lemma 7.1.** *Consider $q_a \in Q$ a solution of Equation 29 and $\gamma \in (0,1)$. The minimization problem*

$$\min_{q\in Q} \left\{ \gamma \tilde{F}(M_\psi(q)) + (1-\gamma)\|q - q_a\|_{W^1} \right\} \tag{32}$$

*is equivalent to the linear problem*

$$\begin{aligned}
\min_{(q,\mu,\eta)} \quad & \gamma \tilde{F}(M_\psi(q)) + (1-\gamma) \sum_{(j,s)\in I} (\mu_j^s + \eta_j^s) \\
s.\ t. \quad & q \in Q, \\
& \mu_j^s - \langle q, \omega_j^s\rangle \geq -\langle q_a, \omega_j^s\rangle, \quad \forall (j,s) \in I, \\
& \eta_j^s + \langle q, \omega_j^s\rangle \geq \langle q_a, \omega_j^s\rangle, \quad \forall (j,s) \in I, \\
& \mu_j^s \geq 0, \quad \eta_j^s \geq 0, \quad \forall (j,s) \in I.
\end{aligned} \tag{33}$$

**Proof.** Let $(\hat{q}, \hat{\mu}, \hat{\eta})$ be a solution of Problem 33. So, $\hat{\mu}$ must satisfy

$$\hat{\mu}_j^s = \max\{\langle \hat{q} - q_a, \omega_j^s \rangle, 0\}, \quad \forall (j, s) \in I,$$

because if we fix $\hat{q}$ we are minimizing $\mu_j^s$ with the constraints $\mu_j^s \geq 0$ and $\mu_j^s \geq \langle \hat{q} - q_a, \omega_j^s \rangle$. The same argument leads to

$$\hat{\eta}_j^s = \max\{-\langle \hat{q} - q_a, \omega_j^s \rangle, 0\}, \quad \forall (j, s) \in I.$$

Now, suppose that $q^*$ is a solution of Problem 32 but not a solution of Problem 33. Then,

$$\min_q \left\{ \gamma F(M_\psi(q)) + (1 - \gamma) \|q - q_a\|_{W^1} \right\} = \gamma F(M_\psi(q^*)) + (1 - \gamma) \|q^* - q_a\|_{W^1}$$

$$= \gamma F(M_\psi(q^*)) + (1 - \gamma) \sum_{(j,s) \in I} |\langle \hat{q}^* - q_a, \omega_j^s \rangle|$$

$$= \gamma F(M_\psi(q^*)) + (1 - \gamma) \sum_{(j,s) \in I} \left( \langle q^* - q_a, \omega_j^s \rangle^+ + \langle q^* - q_a, \omega_j^s \rangle^- \right).$$

If we define $(\mu^*, \eta^*)$ as

$$(\mu^*)_j^s := \max\{\langle q^* - q_a, \omega_j^s \rangle, 0\}, \quad \forall (j, s) \in I$$

and

$$(\eta^*)_j^s := \max\{-\langle q^* - q_a, \omega_j^s \rangle, 0\}, \quad \forall (j, s) \in I,$$

we have that $(q^*, \mu^*, \eta^*)$ satisfies the constraints of Problem 33. Hence, we can write

$$\min_q \left\{ \gamma F(M_\psi(q)) + (1 - \gamma) \|q - q_a\|_{W^1} \right\} = \gamma F(M_\psi(q^*)) + (1 - \gamma) \sum_{(j,s) \in I} \left( (\mu^*)_j^s + (\eta^*)_j^s \right)$$

$$> \gamma F(M_\psi(\hat{q})) + (1 - \gamma) \sum_{(j,s) \in I} \left( \hat{\mu}_j^s + \hat{\eta}_j^s \right) = \gamma F(M_\psi(\hat{q})) + (1 - \gamma) \sum_{(j,s) \in I} \left( \langle \hat{q} - q_a, \omega_j^s \rangle^+ + \langle \hat{q} - q_a, \omega_j^s \rangle^- \right)$$

$$= \gamma F(M_\psi(\hat{q})) + (1 - \gamma) \|\hat{q} - q_a\|_{W^1}.$$

This is a contradiction. A very similar argument proves that if $(\hat{q}, \hat{\mu}, \hat{\eta})$ is a solution of Problem 33, $\hat{q}$ must be solution of Problem 32. ■

**Remark 7.3.** *Instead of $\tilde{\Sigma}$ we can use a different covariance matrix. For example, if we do not have information on the intra-day price we can just use the covariance matrix of $\psi$.*

*Nevertheless, it is recommended to use the covariance matrix* $\tilde{\Sigma} = \Sigma(\psi) + \Delta(\psi, \delta)$, *because* $\tilde{\Sigma}$ *is a positive define matrix, i.e.* $\langle q, \tilde{\Sigma} q\rangle > 0$ *for all* $q \in Q$.

Thus we are led to the following algorithm:

**Algorithm 7.1.**

1. *Solve*
$$\min_{q\in Q}\langle q, \tilde{\Sigma} q\rangle$$
*and choose a solution* $q_a$.
2. *Perform a Cholesky decomposition of* $\tilde{\Sigma} = WW'$.
3. *Solve*
$$\min_{q\in Q}\Big\{\gamma\tilde{F}(M_\psi(q)) + (1-\gamma)\|q - q_a\|_{W^1}\Big\}.$$

*7.4. Numerical Issues*

In practice, solving this two-step problem does not add a significant complexity to the problem of minimizing $\tilde{F}(M_\psi(q))$. We confirm in Table 4.1 that the ESL Model has $O(N_s)$ variables and $O(N_s)$ constraints. The problem 26 is a linear programming problem with $N_aT$ more variables and $4N_aT$ more constraints. However $N_s \gg TN_a$, and thus the operational time of the optimization with regularization will be of the same order of the optimization without it. Additionally, despite the fact that we are solving a quadratic model before Problem 26, the operational time of the whole model will not be affected because this time for the Variance Model is significantly less than the that ESL Model.

*7.5. Regularization in Optimal Allocation*

In Section 3 we reformulated the allocation models to use them to reflect liquidation strategies. Now, we reformulate the liquidation strategy with the regularization term to use it in the allocation context. There are several alternatives to deal with allocation when we have a multi-objective optimization problem, see [7, 24, 4, 3]. Although, it will depend on the investor profile which is better for his purposes.
The approach that we are going to present takes into account our principle of keeping controlled the operational costs. Thus, consider a linear formulation $\tilde{F}$ of the ESL Model. We assume that we have an initial investment of $I_0$ and that we want to compose a portfolio with the restriction of having an expected value of at least $\varepsilon$ of the maximum value. However, minimizing the risk $\tilde{F}$ and the variance at the same time. To simplify we are going to suppose that $T = 1$. Therefore, we present a formulation that proceeds as follow,

**Algorithm 7.2.** 1. *Solve*

$$\begin{aligned} \max_q \quad & \langle \mathbb{E}(\psi), q\rangle \\ s.\ t. \quad & q \geq 0, \\ & \sum_{i=1}^{N_a} q_i m_i \varphi_i^0 = I_0. \end{aligned} \tag{34}$$

*and choose a solution* $q_e$.

2. *Solve*

$$\begin{aligned} \min_q \quad & \langle q, \tilde{\Sigma} q\rangle \\ s.\ t. \quad & q \geq 0, \\ & \langle \mathbb{E}(\psi), q\rangle \geq \varepsilon \langle \mathbb{E}(\psi), q_e\rangle, \\ & \sum_{i=1}^{N_a} q_i m_i \varphi_i^0 = I_0. \end{aligned} \tag{35}$$

*and choose a solution* $q_a$.

3. *Perform a Cholesky decomposition of* $\tilde{\Sigma} = WW'$.
4. *Solve*

$$\begin{aligned} \min_q \quad & \left\{ \gamma \tilde{F}(M_\psi(q)) + (1-\gamma)\|q - q_a\|_{W^1} \right\} \\ s.\ t. \quad & q \geq 0, \\ & \langle \mathbb{E}(\psi), q\rangle \geq \varepsilon \langle \mathbb{E}(\psi), q_e\rangle, \\ & \sum_{i=1}^{N_a} q_i m_i \varphi_i^0 = I_0. \end{aligned} \tag{36}$$

In conclusion, there are several applications that can be given to the $\|\cdot\|_{W^1}$, depending on the focus of the problem. For example, we can use it as a constraint if we wanted to impose a maximum value of variance. Also, we can use it in the context of allocation problems.

## 8. Illustrative Examples

This section presents a few examples that illustrate the claims of Section 7. The first one in Section 8.1 concerns the same portfolio used in Section 5 but using real data of the underlying asset; we are going to study the effect of using $\|\cdot\|_{W^1}$ for different covariance matrices. The second one in Section 8.2, wherein in a new portfolio we will remove the restriction of sales by day and also consider the operational implications of using the $\|\cdot\|_{W^2}$ semi-norm instead of $\|\cdot\|_{W^1}$. Also, we will use the $\|\cdot\|_{W^1}$ semi-norm as a constraint rather than as an objective function.


Figure 7: Daily Prices for SCTY.

The upshot is that the use of regularization in the objective function can help improve the robustness of the liquidation strategies without significant increase in the complexity. Furthermore, this is done while still keeping the financial interpretation and relevance of the model.

*8.1. Changing the Covariance Matrix*

Consider the same portfolio like the one illustrated in Section 5. However, instead of using a fictitious asset, we shall use the SolarCity Corp (SCTY)[2] share as the underlying asset. Hence, we fit an ARMA-GARCH model (see [6]) to the log-returns of the historical data. Figure 7 shows the historical prices of SCTY and Figure 8 shows the histogram of the log-return. Then, we simulate $500,000$ scenarios.

To simplify the analysis, in this example we will not run the Simpleminded Model. For the Variance Model, we use all the scenarios to calculate the covariance matrix. Also, motivated by the robustness of the Example in Section 5, we use $7,000$ scenarios in the ESL Model. Finally, in the ESL Model, we use the confidence level of $\beta = 0.05$[3].

Table 4 shows the statistics for $M_\psi(q)$. There we can see that the results are similar to the results of Example 5. Indeed, the Variance and ESL models have the best results in general.

Moreover, let us add a perturbation of $\varphi$ as in Section 6, i.e., instead of using $\psi$ we used $\tilde{\psi}^t = \psi^t(1+\delta^t) + m\varphi^0\delta^t$ and simulated the perturbation $\delta$ using the intra-day information

[2]Prices were provided by the TradeStation Academic Program through the TradeStation platform.
[3]The optimization problems were run using Gurobi Optimizer, see [1].

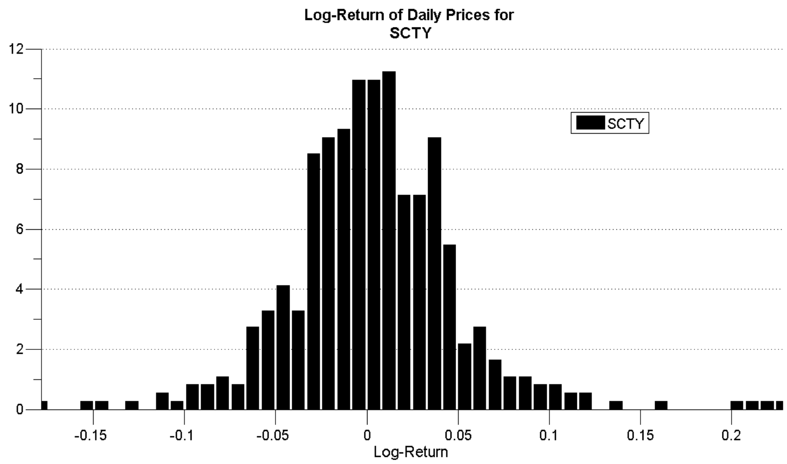


Figure 8: Log Returns for SCTY.

| *Models* | *Std. Des.* | *Min* | $CVaR_{0.05}$ | $VaR_{0.05}$ |
|---|---|---|---|---|
| Variance | 1,379 | 42,240 | 5,710 | 3,879 |
| Ex. Sh. Loss | 1,526 | 44,290 | 5,639 | 3,874 |

Table 4: Statistics of $M_{\psi}(q)$ for Example 8.1.

| *Models* | *Std. Des.* | *Min* | $CVaR_{0.05}$ | $VaR_{0.05}$ |
|---|---|---|---|---|
| Variance | 1,407 | 45,540 | 5,744 | 3,914 |
| Ex. Sh. Loss | 1,562 | 47,690 | 5,698 | 3,903 |

Table 5: Statistics of $M_{\tilde{\psi}}(q)$ for Example 8.1.

| *Models* | *Std. Des.* | *Min* | $CVaR_{0.05}$ | $VaR_{0.05}$ |
|---|---|---|---|---|
| Ex. Sh. Loss $\gamma = 0.90$ | 1,612 | 47,800 | 5,953 | 3,997 |

Table 6: Statistics of $M_{\tilde{\psi}}(q)$ with $\|\cdot\|_{W1}$ for Ex. 8.1, with $W$ s. t. $\Sigma = WW'$.

| *Models* | *Std. Des.* | *Min* | $CVaR_{0.05}$ | $VaR_{0.05}$ |
|---|---|---|---|---|
| Ex. Sh. Loss $\gamma = 0.90$ | 1,412 | 46,900 | 5,674 | 3,881 |

Table 7: Statistics of $M_{\tilde{\psi}}(q)$ with $\|\cdot\|_{W1}$ for Ex. 8.1, with $W$ s. t. $\Sigma + \Delta = WW'$.

of SCTY. The intra-day prices were taken with intervals of one minute during 60 days, and we fitted a non-parametric distribution (see [7]). So, Table 5 shows the statistics of $M_{\tilde{\psi}}(q)$ for the different models. Although the risks do not change significantly, they get worse when we compared it with the results of the same model in Table 4.
Furthermore, we ran the ESL model adding the regularization term $\|\cdot\|_{W^1}$ as in Equation 32 using a Cholesky decomposition. We do this for two covariance matrices $\Sigma(\psi)$, $\Sigma(\psi) + \Delta(\psi, \delta)$, the results are displayed in Table 6 and Table 7, respectively. We can infer that $\Sigma(\psi) + \Delta(\psi, \delta)$ is the appropriate covariance matrix to use because it reduces almost all the risk factors as compared with Table 5. Indeed, when we use $\Sigma$, we are not adding any information to the intra-day risk, we are just controlling the variance.

In the case of the ESL Model, as we saw in Section 2, it seems to be very stable. Nevertheless adding the regularization, especially $\Sigma+\Delta$, it helps to improve the robustness (see Figure 10) and reduce the risk (CVaR) without increasing the operational time as we show in Figure 9.

### *8.2. An Example without Daily Limitation*

Now we have a simpler portfolio using the same underlying asset STCY of Example 8.1 however we eliminate the restrictions for the maximum we can sell or buy per day. The purpose of this example is to see the effect of the regularization when we did not use a daily limit to the sale or purchase of the assets. Also, we will analyze the problem of adding the quadratic term as regularization as Equation 31 instead of linear regularization. As a reference, we show in Table 9 the results of the Simpleminded Model and the Variance Model. For an additional analysis, we add the expected value of the loss of $M_{\tilde{\psi}}(q)$, $\hat{\mathbb{E}}(M^-_{\tilde{\psi}}(q))$, in all statistical tables. Also, from now on, we are going to use only the regularization term in $\tilde{\Sigma} = \Sigma + \Delta$. We ran the ESL Model for different values of $\gamma$ between 0.9 and


Figure 9: ESL Model for Example 8.1. Left: CVaR of the loss of $M_{\tilde{\psi}}(q)$. Right: Operational Time.

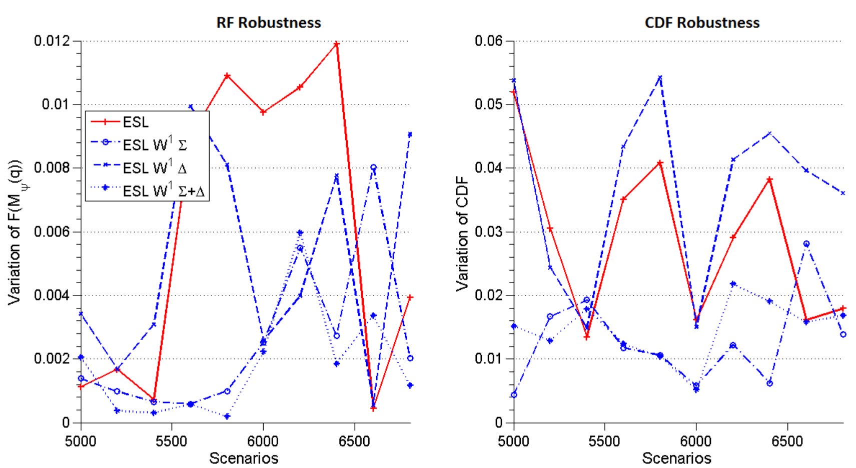


Figure 10: Robustness of ESL Model for Example 8.1. Left: RF Robustness. Right: CDF Robustness.

| *Asset* | *Product* | *Position* | *Exp Day* | *Strike* | *Max p/ Day* | *Initial Day* |
|---|---|---|---|---|---|---|
| 1 | Option Call | -2200 | 60 | 70 | N/A | 5 |
| 2 | Option Put | 2000 | 60 | 70 | N/A | 5 |
| 3 | Forward | 2000 | 60 | 70 | N/A | 2 |

Table 8: Portfolio for Example 8.2.

| *Models* | *Std. Des.* | *Min* | $CVaR_{0.05}$ | $VaR_{0.05}$ | $\hat{\mathbb{E}}(X^-)$ |
|---|---|---|---|---|---|
| Simpleminded | 15,938 | 262,680 | 39,686 | 29,453 | 13,674 |
| Variance | 1,182 | 25,550 | 3,515 | 2,522 | 1,098 |

Table 9: Statistics of $M_{\tilde{\psi}}(q)$ for Ex. 8.2.

1.0, with a sample size of $7,000$. Remembering that $\gamma = 1.0$ means that the model does not have regularization. In Table 10 we observe the effect of adding the regularization.

| *Models* | *Std. Des.* | *Min* | $CVaR_{0.05}$ | $VaR_{0.05}$ | $\hat{\mathbb{E}}(X^-)$ |
|---|---|---|---|---|---|
| ESL $\gamma = 1.00$ | 1,618 | 39,181 | 3,493 | 2,462 | 1,214 |
| ESL $\gamma = 0.99$ | 1,559 | 34,945 | 3,420 | 2,437 | 1,211 |
| ESL $\gamma = 0.98$ | 1,514 | 39,659 | 3,476 | 2,424 | 1,168 |
| ESL $\gamma = 0.97$ | 1,491 | 36,373 | 3,426 | 2,392 | 1,125 |
| ESL $\gamma = 0.96$ | 1,245 | 30,786 | 3,459 | 2,460 | 1,048 |
| ESL $\gamma = 0.95$ | 1,224 | 29,744 | 3,470 | 2,466 | 1,053 |
| ESL $\gamma = 0.94$ | 1,258 | 31,595 | 3,464 | 2,447 | 1,037 |
| ESL $\gamma = 0.93$ | 1,229 | 29,193 | 3,477 | 2,468 | 1,048 |
| ESL $\gamma = 0.92$ | 1,193 | 28,301 | 3,497 | 2,483 | 1,078 |
| ESL $\gamma = 0.91$ | 1,183 | 25,612 | 3,506 | 2,525 | 1,096 |
| ESL $\gamma = 0.90$ | 1,182 | 25,692 | 3,513 | 2,516 | 1,095 |

Table 10: ESL Model. Statistics of $M_{\tilde{\psi}}(q)$ for ESL with $\|\cdot\|_{W1}$ for Ex. 8.2.

By reducing the value of $\gamma$ until 0.97, we reduce four of the five risk measures including the Expected Shortfall (CVaR). This reduction is caused by the objective function $\gamma F_{ESL}(M_\psi(q)) + (1-\gamma)\|q - q_a\|_{W1}$. On one side it seeks to minimize the expected shortfall of the loss of $\langle \psi, q \rangle$ that does not see the intra-day variations. On the other one, the term on the right hand side minimize the variance of $\langle \tilde{\psi}, q \rangle$, hence to control (not minimize) its expected shortfall. Therefore, the combination of minimizing the Expected Shortfall of $M_\psi(q)$ and controlling the Expected Shortfall of $M_{\tilde{\psi}}(q)$ produces a strategy that is better

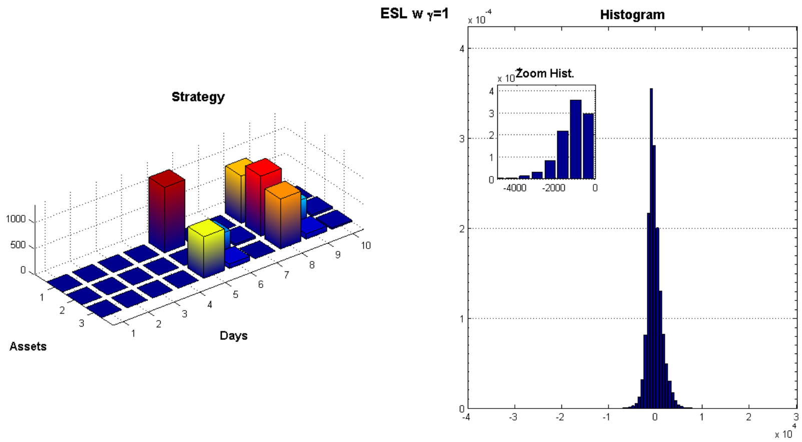


Figure 11: Strategy $q$ and distribution of $M_{\tilde{\psi}}(q)$ for the ESL Model for Ex. 8.2.

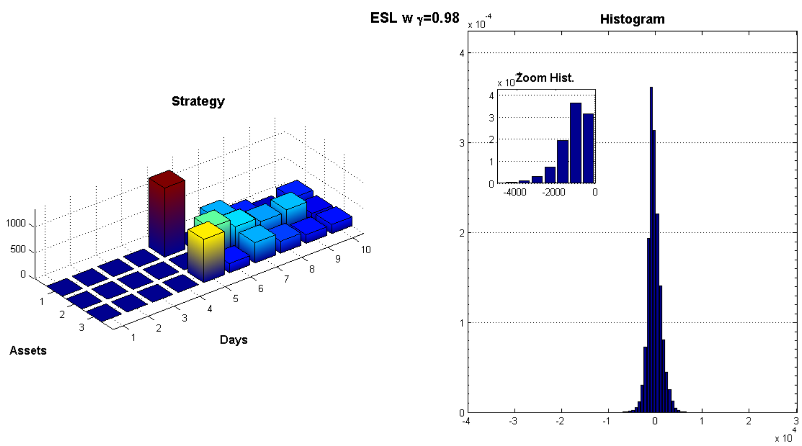


Figure 12: Strategy $q$ and distribution of $M_{\tilde{\psi}}(q)$ for the ESL Model with $\gamma = 0.8$ for Ex. 8.2.

than only minimizing the Expected Shortfall.

Figures 11 and 12 show the difference of using the regularization term. We can see that using just $\gamma = 0.98$ the solution is better distributed in time giving an improvement of the risk measure.

To study the robustness, we will also consider the model with the quadratic regularization, i. e.,

$$\gamma F_{ESL}(M_\psi(q)) + (1-\gamma)\|q - q_a\|^2_{W^2}.$$

Thus, we run the three ESL models, without regularization, with $\|\cdot\|_{W^1}$ regularization and with $\|\cdot\|_{W^2}$ regularization for $5,000$ to $7,000$ scenarios increasing by 200. In all those, we set $\gamma = 0.9$.

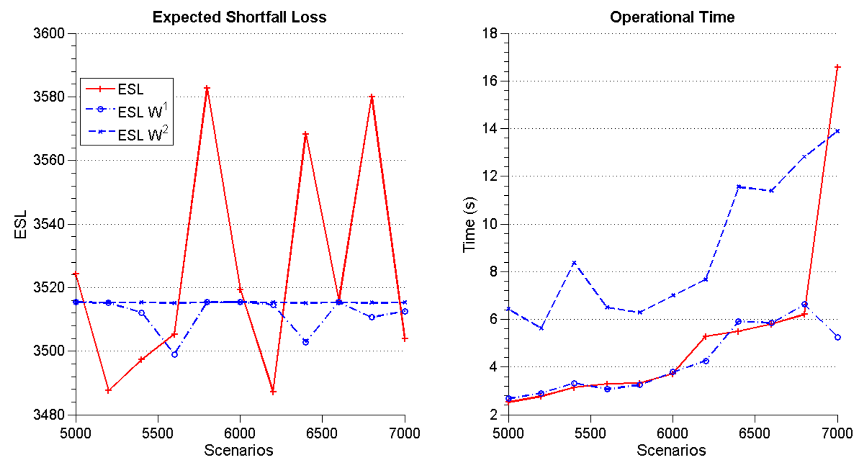


Figure 13: ESL Model for Example 8.2. Left: CVaR of the loss of $M_{\tilde{\psi}}(q)$. Right: Operational Time.

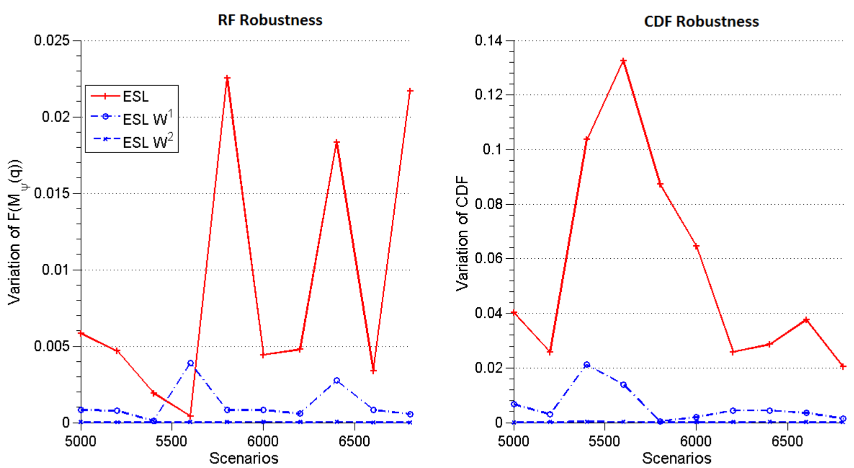


Figure 14: Robustness of ESL Model for Example 8.2.Left: RF Robustness. Right: CDF Robustness.

In Figures 13 and 14 it looks like using the linear term we can reduce the ESL and improve the stability without affecting the runtime. On the other hand, the fact that with $\|\cdot\|^2_{W^2}$ the model is extremely stable is due to the fact the quadratic norm weighs too much close to compared with the $ESL$, hence the optimization leads to using only the variance. Moreover, using the quadratic term takes significantly longer than not using the regularization or using it with $\|\cdot\|_{W^1}$. Consequently, the results in Table 11 confirm our conclusions, namely, that using $\|\cdot\|^2_{W^2}$ the solution is almost exactly the solution for variance (see Table 9). However, using $\|\cdot\|_{W^1}$ keeps the standard deviation close to the minimum value but also reduce the conditional VaR.

| *Models* | *Std. Des.* | *Min* | $CVaR_{0.05}$ | $VaR_{0.05}$ | $\hat{\mathbb{E}}(X^-)$ |
|---|---|---|---|---|---|
| ESL | 1,634 | 44,638 | 3,579 | 2,482 | 1,212 |
| ESL w $W^1$, $\gamma = 0.9$ | 1,182 | 25,830 | 3,510 | 2,514 | 1,095 |
| ESL w $W^2$, $\gamma = 0.9$ | 1,182 | 25,571 | 3,515 | 2,521 | 1,097 |

Table 11: ESL Model. Statistics of $M_{\tilde{\psi}}(q)$ for ESL in Ex. 8.2.

## 9. Conclusions

We have reviewed two of the most important risk measures and their optimization models. We introduced a technique which incorporates the advantages of the Variance Model to this linear model. We did this by performing a Cholesky decomposition of the covariance matrix and adding to a Tikhonov regularization term. This regularization term is used as an $L^1$ semi-norm which added to the linear model has several practical properties. Indeed, we can control the price perturbation caused by the impact on the price as the result of the liquidation process or the intra-day variations, improve the robustness and control the variance. All of these without further computational cost.


## Acknowledgments

We thank Claudia Sagastizabal for a number o very fruitful discussions and suggestions. JPZ was supported by CNPq, FAPERJ, and the Pensa-Rio program of FAPERJ.


## Bibliography


[1] Gurobi Optimization, Inc., Gurobi Optimizer Reference Manual, 2014, www.gurobi.com

[2] CME Group, CME CORE: Clearing Online Risk Engine, www.cmegroup.com/clearing/risk-management

[3] Y. Elahi and M. Ismail Abd Aziz, Mean-Variance-CvaR Model of Multi-portfolio Optimization via Linear Weighted Sum Method, 2014, Mathematical Problems in Engineering, 2014, 104064

[4] R. Aboulaich, R. Ellaia and S. El Moumen, The Mean-Variance-CVaR model for Portfolio Optimization Modeling using a Multi-Objective Approach Based on a Hybrid Method, 2010, 10.1051/mmnp/20105717, Math. Model. Nat. Phenom., 103-108

[5] A. Tikhonov, A. Goncharsky, V. Stepanov and A. Yagola, Numerical Methods for the Solution of Ill-Posed Problems, 978-90-481-4583-6, Springer Science+Business Media Dordrecht, Mathematics and Its Applications, 1995 Moscow, Russia

[6] New York, USA, Terence C. Millis, 978-0-521-88381-8, Cambridge University Press, The Econometric Modelling of Financial Time Series, 2012

[7] Berlin, Meucci, Attilio, 978-3-540-22213-2; 3-540-22213-8, Springer-Verlag, Springer Finance, Risk and asset allocation, 2005

[8] A. Izmailov, M. Solodov, IMPA, IMPA, Otimização - volume 1, 2005

[9] S. Boyd and L. Vandenberghe, Convex Optimization, 2004, Cambridge University Press

[10] M. Avellaneda and R. Cont, Close-Out Risk Evaluation (CORE): A New Risk Management Approach for Central Counterparties, 2013, SSRN eLibrary, ssrn.com/paper=2247493

[11] L. Vicente, F. Cerezetti, S. De Faria, T. Iwashita, O.R. Pereira, Managing risk in multi-asset class, multi-market central counter-parties: The CORE approach, 2014, Journal of Banking and Finance, 10.1016/j.jbankfin.2014.08.016

[12] R. Almgren and N. Chriss, Optimal Execution of Portfolio Transactions, 2000, Journal of Risk, 3, 539

[13] M. Kapsos, S. Zymler, N. Christofides and B. Rustem, Optimizing Omega Ratio Using Linear Programming, 2012, Journal of Computational Finance

[14] M. Schweizer, Mean-variance hedging for general claims., 1992, Annals of Applied Probability, 2, 171-179

[15] R. Cont, R. Deguest and G. Scandolo, Robustness and Sensitivity Analysis of Risk Measurement Procedures, 2010, Quantitative Finance, 10, 6, 593606

[16] K. van den Doel, U. Ascher and E. Haber, The lost honour of L2 based regularization, 2013, Large Scale Inverse Problems, 181-203

[17] 1998, Journal of Financial Markets, 1, Optimal Control of Execution Costs, D. Bertsimas and A. W. Lo, 1-50

[18] 2013, Statistic & Risk Modeling, 30, Loss-based Risk Measures, R. Cont, R. Deguest and X. Dong He, 133-167

[19] 2002, The Finance Development Centre, A Universal Performance Measure, C. Keating, W. F. Shadwick

[20] 2002, ISE Dept., Univ. of Florida, 7, Deviation measures in risk analysis and optimization, R. T. Rockafellar, S. Uryasev and M. Zabarankin

[21] 1998, Management Science, 44, 5, A Minimax Portfolio Selection Rule with Linear Programming Solution, M. R. Young, 673-683

[22] 1966, Journal of Business, 39, Mutual Fund Performance, W. F. Sharpe., 119-138

[23] 2000, Journal of Risk, 2, Optimization of Conditional Value-at-Risk, R. T. Rockafellar and S. Uryasev, 21-41

[24] 1952, Journal Finance, 7, Portfolio Selection, H. Markowitz, 77-91

[25] Black, Fisher and Scholes, Myron, Journal of Political Economy, 637–659, The Pricing of Options and Corporate Liabilities, 81, 1973

[26] New York, Karatzas, I. and Shreve, S. E., Second, xxiv+470, Springer-Verlag, Graduate Texts in Mathematics, Brownian motion and stochastic calculus, 113, 1991

[27] Artzner, Philippe and Delbaen, Freddy and Eber, Jean-Marc and Heath, David, Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics, 0960-1627, Math. Finance, 91B30 (91B28), MR1850791 (2002d:91056), Yanhong Wu, 3, 203–228, Coherent measures of risk 9, 1999

[28] Acerbi, C., Journal of Banking and Finance, 1505–1518, Spectral Measures of Risk: A coherent representation of subjective risk aversion, 26, 2002

[29] Frittelli, Marco and Scandolo, Giacomo, Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics, 0960-1627, Math. Finance, 91B30 (60A10 91B70), MR2253119 (2007g:91078), Jostein Paulsen, 4, 589–612, Risk measures and capital requirements for processes, 16, 2006